THIS REPORT HAS BEEN DELIMITED
AND CLEARED FOR PUBLIC RELEASE
UNDER DOD DIRECTIVE 5200.20 AND
NO RESTRICTIONS ARE IMPOSED UPON
ITS USE AND DISCLOSURE.

DISTRIBUTION STATEMENT A

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION UNLIMITED.

RCDL-C 23

ATI 308

UNCLASSIFIED

STI ATI No. 205 308 / ASTIA FILE COPY

TO:
...E CENTER
...NGRESS
...25, D.C.

...SITY PSYCHOLOGICAL RESEARCH CENTER

195...

A FACTORIAL ANALYSIS OF DEPTH PERCEPTION TESTS

LYLE H. LANIER and MARGARET E. TRESSELT

...ARY 1953

This study was [illegible] out under Contract No. WI.W-2495 between New York Univ[illegible] and the Department of the Army, Personnel Research Section [illegible] Adjutant General's Office.

A FACTORIAL ANALYSIS OF DEPTH PERCEPTION TESTS

Table of Contents

A FACTORIAL ANALYSIS OF DEPTH PERCEPTION TESTS[1]

I. THE PROBLEM

1. Background

This study is partly a sequel to the investigation of visual acuity conducted by the Personnel Research Section, AGO, Department of the Army (1). During the past war, dissatisfaction with tests of "visual acuity" led to efforts by the Army-Navy-National Research Council Vision Committee to standardize visual examinations, in the hope that uniform results might be secured throughout the armed forces. In the course of this work the question arose as to whether or not the wide variety of tests in use were actually measuring a single homogeneous function. It is of great practical as well as theoretical significance to know whether several relatively independent aspects of visual discrimination are involved in the various measures of "visual acuity."

The results of the AGO factorial analysis showed clearly that "visual acuity" as measured by different tests is not a simple, unidimensional variable. Instead, several essentially unrelated aspects of visual discrimination apparently contribute to the variance of most of the measures. Tests superficially alike differed markedly in "factorial" composition. A similar situation was found with tests of depth perception and of phorias, which were analyzed in a second part of the research. The following general conclusion was reached:

> "Since all presently available acuity and depth vision tests, wall chart or machine, have large parts of their variance determined by non-pertinent factors, much remains to be done to make them really satisfactory.... Any validity studies based upon a single type of visual test can only with great danger be used to predict action based upon some other method supposedly measuring the same function" (1, p. 140).

The depth perception tests used in the AGO analysis were all based upon binocular disparity in the views of test objects presented by means of stereoscopic optical devices. The four depth variables were (a) the "far depth" measures of the American Optical Company's Sight-Screener, the Bausch and Lomb Ortho-Rater and the Keystone Telebinocular, and (b) the "near depth" score of the Sight-Screener. All of them were supposedly measuring stereoscopic acuity. Yet the correlations among the measures secured with different machines were only moderately high, ranging from .31 to .63. The factor called "Depth Perception" showed significant

[1]The following graduate students served as research assistants throughout the period of this study: Paul Bakan, Benjamin H. Brown and Sidney Weinstein. Their unfailing cooperation is gratefully acknowledged.

loadings on all of these depth variables; but the variance of the scores for some of the tests was due as much to the general acuity factor called "Retinal Resolution" as to "Depth Perception." In all of the depth tests the "specific and error" variance was high.

With such a high degree of specificity among these superficially similar tests of apparent depth perception, it is not surprising that their correlation with tests involving real depth would be low. Such is the trend among the correlations reported. Fonda et al. (2) found virtually no correlation between Howard-Dolman scores and the Sight-Screener depth test, and scores on the Howard-Dolman rod test showed quite low correlations with Ortho-Rater scores, according to Imus (11). Using subjects with 20/20 vision, he found correlations ranging from .14 to .23 between accuracy scores on the Howard-Dolman and the Ortho-Rater depth test. The correlations between Howard-Dolman variability scores and the Ortho-Rater scores range from .20 to .40. Moffie (13) found higher correlations between Howard-Dolman and Ortho-Rater scores: .49 and .54 for accuracy and variability scores on the Howard-Dolman, respectively. Moffie also used subjects with 20/20 visual acuity or better. Such a restriction of the range of subjects in visual acuity should yield lower correlations than would be found with an unselected sample, if performance on the depth tests is influenced by visual acuity. This inference is supported by the results of the AGO study showing the influence of visual acuity upon depth perception. Experimental studies (e.g., 9) have likewise shown a high correlation between visual acuity and stereo acuity.

In view of the moderate size of the intercorrelations among binocular tests of apparent depth, it might perhaps be predicted that different binocular tests of real depth perception would likewise fail to show a high degree of correspondence. One study (15) reports a correlation of .43 between the Howard-Dolman test and the Verhoeff Stereopter (18). The two visual situations presented by these two tests are alike in depending upon binocular parallax, but the tests are quite different in other respects, including the mode of response of the subject. Evidently the differences contribute more to the determination of performances in the two tests than does the "communality."

All of the foregoing studies have been concerned with binocular depth perception, which arises from the disparity of the two views of tridimensional objects presented to the two eyes. But perception of depth and distance is possible with one eye and the "monocular" cues are usually operative concurrently with the binocular. The important monocular factors include relative size, interposition, linear perspective, aerial perspective, relative rate of movement, light and shade, and accommodation (6). These conditions become increasingly important as the discriminations involved occur at greater and greater distances from the subject. Since the eyes are only about two and a half inches apart, they will have virtually the same view of objects 100 yards away. Hirsch and Weymouth (8) report that at 122 meters binocular distance perception is twice as good as monocular, and estimate that at about 1200 meters the two types would be approximately equal in effectiveness. But these figures do not necessarily mean that binocular disparity (stereo acuity) produces all of the superiority of the two eyes over one. The coordinated use of monocular cues by both eyes is probably also an important condition of depth perception at such distances. It seems likely that the depth and distance discriminations involved in many of the tasks of the flyer and of the motor vehicle operator are determined mainly by monocular cues. Moreover, the numerous and complex monocular cues may operate differentially at different distances from the observer.

This general conclusion is supported by Kirschberg's (12) report of the absence of relationship between flying ability and performance on the tests of binocular depth perception.

There has been a comparative dearth of research on monocular depth perception, which is explained by Gibson (5) in terms of the failure to identify and control the stimulus variables in question. In the case of binocular depth perception the development of stereoscopic and associated instruments permit of quantitative control of the stimulus variable of retinal disparity, with a consequent emphasis upon research in stereo acuity. Practical impetus to research on stereo acuity was also provided in World War II by the necessity to select and train operators of stereoscopic range finders and height finders. No comparable interest or research effort occurred for monocular depth perception. Gibson's development of a size estimation test for the Army Air Forces (6) is an important exception to this trend. And it should be noted that a considerable amount of important basic research has been done on two aspects of monocular depth perception. Graham and his associates (7, 19) have established significant quantitative relationships for the factors which influence monocular movement parallax thresholds. And a considerable amount of work has been done on the relationship between size and distance in depth perception (4, 10).

2. The Present Study

In the light of previous research, "depth perception" is apparently not a homogeneous behavior variable but instead is a broad class name for a variety of visual discriminations. Varying degrees of "communality" seem to exist among these diverse aspects of depth and distance discrimination. It is important to know more about the "structural" interrelationships among these several dimensions of depth perception, as well as to know how discrimination of distance is related to other visual functions. The present work attempts to study these relationships. Such knowledge should have both theoretical and practical value. It should contribute towards an integrated science of visual behavior by providing clues to the nature of the dependent variables involved. And it should provide guidance to individuals and agencies concerned with the selection, training and management of men who must judge depth and distance in the course of occupational duties.

More specifically, this study has the following objectives:

(a) To secure norms of performance for a fairly large group of subjects in a variety of visual test situations, including both monocular and binocular depth discriminations. The tests would be given twice, with an interval of about one week between the two series.

(b) To study the interrelationships among performances on the tests by means of factor analysis, with particular reference to the determinants of variance in depth perception tests.

(c) To establish the reliability both of individual tests and of the factorial analysis, by comparison of results secured in the two series of tests.

(d) To determine the possible effects of practice upon the different types of visual behavior under study.

II. THE TEST VARIABLES

The selection of the visual variables was dictated mainly by the following considerations:

1. The test situations would include both non-depth and depth variables.

2. Both monocular and binocular depth variables would be included.

3. The binocular tests would involve both "real" depth and simulated stereoscopic depth.

4. The monocular tests would be as varied as possible, consistent with the scope of the study and the technical feasibility of control of conditions.

5. The non-depth variables would include both acuity and phoria measurements comparable with those made in the earlier AGO study.

6. The series of tests would be administered within a single period of not longer than two hours - to be repeated at a later period, preferably one week later.

7. In order to include a wide variety of test situations within the limited period, each component was made as short as seemed consistent with a reasonably high level of reliability (determined in exploratory experimentation).

Twenty-eight visual variables were finally included in the study. Together with the age of the subject, they are listed in Table I in the order of administration to the subjects. The following is a classification of these variables, the numbers in parentheses referring to the numbers in Table I:

TABLE I

List of Variables Included in the Study

1. Age -- the ages of the subjects (all male) ranged from 17 to 39 years.
2. Howard-Dolman binocular depth perception test (at 20 feet)
3. Howard-Dolman adaptation for monocular motion parallax
4. Sight-Screener -- far stereopsis (Sloan's circles)
5. Sight-Screener -- far acuity (Sloan's Landolt rings - both eyes)
6. Sight-Screener -- near stereopsis (Sloan's circles)
7. Sight-Screener -- near acuity (Sloan's Landolt rings - both eyes)
8. AAF Distance Estimation Test CP212-A (Gibson, size-at-a-distance)
9. Ortho-Rater F-6 -- far stereopsis
10. Ortho-Rater F-5 -- checkerboard: far acuity, left eye
11. Ortho-Rater F-4 -- checkerboard: far acuity, right eye
12. Ortho-Rater F-3 -- checkerboard: far acuity, both eyes
13. Ortho-Rater N-3 -- checkerboard: near acuity, left eye
14. Ortho-Rater N-2 -- checkerboard: near acuity, right eye
15. Ortho-Rater N-1 -- checkerboard: near acuity, both eyes
16. Ortho-Rater F-1 -- far vertical phoria
17. Ortho-Rater F-2 -- far lateral phoria
18. Ortho-Rater F-7 -- color
19. Ortho-Rater N-4 -- near vertical phoria
20. Ortho-Rater N-5 -- near lateral phoria
21. Telebinocular circles: near acuity, both eyes
22. Telebinocular signposts: far acuity, right eye
23. Telebinocular signposts: far acuity, left eye
24. Telebinocular near stereopsis
25. Telebinocular far stereopsis
26. Howard-Dolman-Weinstein (tilted, colored targets at 10 feet)
27. Ames-Tresselt size-depth adjustment
28. Ames-Bakan interposition illusion
29. Verhoeff Stereopter

Test variables 2 - 29 are listed in order of administration. The numbers at the left are used throughout the report to designate the respective variables.

A. Binocular Depth Variables

1) Real depth (2, 26, 29)
2) Simulated depth (4, 6, 9, 24, 25)

B. Monocular Depth Variables

1) Monocular motion parallax (3)
2) Size-at-a-distance estimation (8)
3) Size-depth illusion (27)
4) Interposition illusion (28)

C. Visual Acuity

1) Far acuity (5, 10, 11, 12, 22, 23)
2) Near acuity (7, 13, 14, 15, 21)

D. Phorias

1) Vertical, far and near (16, 19)
2) Lateral, far and near (17, 20)

E. Color (Ortho-Rater) (18)

Descriptions of these test variables are given in some detail in Appendix A, including the numbers of items (discriminations), and the methods of scoring.

It would have been highly desirable to have had more - and also more varied - situations involving the use of monocular cues. But the necessary research and development has simply not been done with these variables, as Gibson has pointed out (4, 5). Extensive experimental work must precede the development of standardized "tests" based upon these complexly intermingled conditions. It was not possible in the present study to go farther than the four variables listed. And even with these factors the experimental base was not too solidly established. But it seemed worth while to check upon the trend in interrelationships among even so few monocular variables, as well as to determine their relationships with the other visual functions listed.

III. PROCEDURE

The study developed in four stages: (a) review of literature, selection and development of the tests; (b) preliminary experimentation, standardization of conditions, and training of examiners; (c) administration of the test and retest series to 228 subjects; (d) analysis of results and preparation of the report. This section describes principally the general conditions and procedure for the administration of the tests. Appendix A contains a description of the several tests.

1. Test Room.

All of the tests were conducted in a large L-shaped room which was subdivided into appropriate booths and one small dark-room. The booths were arranged around the walls of this room so that the several tests could be easily administered in the order listed in Table I. Shades were drawn in the room and no lights were used except those in the test instruments or situations. The walls were painted a neutral gray, except in the Ames-Tresselt room which was painted black.

Electric fans and a large exhaust fan kept the room fairly comfortable.

2. Preliminary Testing - Training of Examiners.

The period June 1 to July 11, 1949, was devoted to the standardization of the tests and to the training of examiners. Preliminary test-retest reliability coefficients were calculated for most of the variables, as a basis for setting the lengths of the tests. The latter factor in turn determined the number of variables that could be included in the battery. As a rule, some 20 - 30 subjects were used in these preliminary tests.

The examiners were all graduate students chosen for dependability and who seemed to achieve a satisfactory degree of rapport with subjects.[2]

Instructions were prepared for each test, in two parts: (a) directions to be read by the subject, or to the subject by the examiner; (b) directions to the examiner. The examiners were first given the tests, and then practiced the administration of each test under supervision until the procedure was well understood. They then tested several additional subjects before beginning the regular series.

3. Procedure with the Regular Test Series.

The subject came to an office adjacent to the test room for a brief preliminary interview. Relevant items of visual and medical history were noted. He was asked about previous tests of depth perception and about his occupational history. An examiner then conducted him to the test room and the series began. A single examiner gave all of the tests, in order, to a given subject. Subjects who had them were tested with their glasses on. The total time varied from 1½ to 2½ hours (the retest sessions were naturally shorter than the initial ones). Although the series became somewhat fatiguing or tiresome to certain subjects, in general the individual tests seemed to be sufficiently short and varied to render these factors relatively unimportant.

[2]In addition to the research assistants, Messrs. Bakan, Brown and Weinstein, the examiners were: Maxwell E. Becker, Leah Blaustein, Melvin Dudkin, Robert Vineberg, Martin Wagner, Bernard Weiss and Alvin Wolf.

The schedule called for the retesting of the subjects one week after the initial series. But many of them did not return as scheduled, and these were retested as soon afterwards as available.

Because of delay in completing equipment for variable 27 (Ames-Tressolt) and 28 (Ames-Bakan), these tests were not included in the first test series for subjects tested during July 1949 (approximately 100 Ss). But by the time these Ss were retested, these two variables were in use; they were later retested on them separately. So both the test and retest series for these two variables followed variable 29 (Verhoeff) for about half the subjects. In view of the virtually complete lack of correlation among these variables, it is not believed that this irregularity influenced the results appreciably.

IV. SUBJECTS

A total of 228 men served as subjects in the study. They were mainly undergraduates in Washington Square College, New York University, although in order to complete the required number of retests a few students in other schools and a few non-students were used.

The ages of the subjects ranged from 17 to 39, with a mean of 21.86. A frequency distribution of ages is shown in Appendix B, Table B-1. Originally it was planned to restrict the age range to 17-25, but when it became difficult to secure subjects it was decided to extend the range and to treat age as a variable in the correlation matrix. Only 38 of the 228 Ss were older than 25 years.

c. Visual acuity tests. The use of our 228 "unselected" college students makes it possible to compare the three machine tests of visual acuity in terms of comparative difficulty and manner of dispersion of scores. In the AGO study most of the 14 wall chart tests gave negatively skewed and somewhat bimodal distributions (1, pp. 82-83). It could not be determined whether or not "the 'real' distribution of visual abilities tested are piled up in the same way" (p. 82). In the present study the several distributions of acuity measurements differ considerably among themselves, which suggests that such variations are functions of the nature of the test targets and of other associated visual conditions. For example, consider the three acuity distributions for the Telebinocular. The far acuity tests with the "sign-posts" (right and left eye tested separately) show marked negative skewness (Tables B-22 and B-23), with 45 to 50 per cent of the subjects getting maximum scores. The Telebinocular circles test (both eyes) is much more difficult and shows far better discrimination among the subjects (Table B-21). The percentage scales of relative difficulty presented in the Telebinocular test manual gives the same value (105) to scores which are reached, respectively, by 40 and by 6 per cent of our subjects (see Tables B-21 and B-22).

Considerable disparity exists between the Sight-Screener and the Ortho-Rater acuity tests in level of difficulty and in form of dispersion of scores.[3] The Sight-Screener far and near acuity scores show marked negative skewness, and the means of the distributions indicate that the test is considerably easier than the Snellen values suggest (the means are around 20/15 to 20/16). The corresponding Ortho-Rater distributions show much more satisfactory dispersion of scores, as well as means which agree more closely with the Snellen "norm." Tables B-12 and B-15 contain these far and near acuity distributions (both eyes). The degree of negative skewness is considerably less and the means lie at approximately 20/19 on the Snellen scale. The four Ortho-Rater distributions for the two eyes separately show similar characteristics, with the far acuity tests having lower mean scores than the near tests (Tables B-10, B-11, B-13, B-14).

d. Phoria tests. The four phoria distributions for the Ortho-Rater are shown in Tables B-16 (far vertical), B-17 (far lateral), B-19 (near vertical) and B-20 (near lateral).

The general form of all four sets of distributions is reasonably satisfactory. Although not strictly "normal," none of these distributions shows the skewness characteristic of many of the variables described in the preceding paragraphs. In the far vertical phoria test more than 40 per cent of the subjects have the same score of "6," but the frequencies tend to diminish fairly regularly in both directions from this mode. (At the near point the mode is "5.") In the case of the lateral phoria measure-

[3] It should be recalled that the Sight-Screener targets used in this study are not the standard ones, but rather are the Landolt rings on the special slide.

ments, the scale covers a wider range and correspondingly the measurements are more widely dispersed.

The means for the phoria measurements show these subjects to have considerably more right hyperphoria than the naval population studied by Sulzman, Cook and Bartlett (17).[4] They report means of 5.17 and 5.11 for test and retest scores, while our means are 6.24 and 6.25, for far vertical phoria. The means are more alike for the near vertical phoria measurements: their values are 4.50 and 4.59, while ours are 4.89 and 4.92. With respect to lateral phorias, their subjects are slightly more exophoric (or less esophoric) than ours at "far distance"; at the near point the relations are reversed.

2. Comparison of Test and Retest Means: Practice Effects

The means, standard deviations and "critical ratios" for all variables are presented in Table II. These values are based upon the frequency distributions in Appendix B. The score units are explained in Appendix A. For variables 2, 3, 26 and 27 the lower the score the better the performance (S adjusted two objects until they were judged to be the same distance away). In the case of the phoria measurements (variables 16, 17, 19 and 20), the scores are not "performance indices" in the same sense as with all of the other tests. For the remainder of the variables, the score varies directly with level of performance.

a. Practice effects for binocular depth tests. It is apparent from Table II that highly significant practice effects appear upon the repetition of the binocular stereopsis tests (variables 2, 4, 6, 9, 24, 25, 26, 29). The differences between the test and retest means for the eight "binocular" variables are all greater than three times their standard errors. It should be noted that the two sets of scores are rather highly correlated (see the last column in Table II), with the result that small differences become "significant" with an N of 228.

In general, the tests of "simulated" depth (Ortho-Rater, Sight-Screener and Telebinocular) show more "significant" practice effects than do the tests of "real" binocular depth. But this may reflect partly the fact that the method of measuring the latter gave opportunity for more variability than was possible with the optical devices. The magnitude of the differences between the test and retest means can be measured roughly by dividing the difference in each case by, say, the standard deviation of the retest distribution. The quotients can then be converted into units of area under the normal probability curve. When the "raw score" differences between means are transformed into these "standard" units, the simulated depth tests as a group are still found to show greater gains upon retest than the "real" depth tests (.10 vs .06 in "normal curve" units). However, the

---

[4] This was the group whose scores were used in the second part of the A.G.O. study.

TABLE II.

Showing Means, Standard Deviations, "Critical Ratios" and Reliability Coefficients for Test and Retest Scores

	Variables	TEST Mean	TEST S. D.	RETEST Mean	RETEST S. D.	Diff. in Means / S.E. Diff. ("Critical Ratio")*	Test-Retest Correlations
1.	Age	21.86	2.96	--	--	--	--
2.	H-Dolman	35.31	27.72	31.85	25.56	3.01	.79
3.	H-D Motion	35.87	10.73	31.42	9.75	5.71	.34
4.	SS F.Stereopsis	11.48	4.43	12.78	3.58	5.82	.65
5.	SS F. Acuity	42.04	10.45	43.38	9.07	2.85	.74
6.	SS N.Stereopsis	13.05	4.09	13.81	3.12	4.47	.79
7.	SS N. Acuity	44.23	8.05	44.80	6.95	1.51	.72
8.	AAF-Size	41.85	8.38	45.92	9.76	7.02	.54
9.	O-R F. Stereo.	5.13	2.17	5.49	2.31	3.60	.72
10.	O-R F.Acuity (L)	9.76	2.43	9.70	2.61	-.46	.70
11.	O-R F.Acuity (R)	9.90	2.62	9.87	2.58	-.34	.87
12.	O-R F.Acuity (B)	10.48	2.04	10.63	2.16	1.50	.72
13.	O-R N.Acuity (L)	10.28	2.38	10.41	2.02	.93	.59
14.	O-R N.Acuity (R)	10.26	2.27	10.47	1.89	1.91	.70
15.	O-R N.Acuity (B)	10.39	1.65	10.60	1.64	2.10	.59
16.	O-R F.V.Phoria	6.24	1.21	6.25	1.17	.18	.70
17.	O-R F.L.Phoria	7.13	2.79	7.18	2.72	.40	.76
18.	O-R Color	4.80	.97	4.87	.91	1.96	.84
19.	O-R N.V.Phoria	4.89	1.23	4.92	1.29	.38	.72
20.	O-R N.L.Phoria	6.84	3.46	6.50	3.44	-2.43	.78
21.	TB N.Acuity (B)	15.63	3.45	16.29	3.36	2.80	.45
22.	TB F.Acuity (R)	7.73	2.94	8.09	2.71	1.09	.79
23.	TB F.Acuity (L)	8.13	2.71	8.20	2.65	.70	.85
24.	TB N.Stereo.	16.00	6.76	17.80	6.70	5.81	.76
25.	TB F.Stereo.	14.91	6.22	16.83	6.45	5.49	.65
26.	H-Dolman (Tilt)	27.61	20.64	24.27	18.19	3.34	.70
27.	Ames (Size)	3.49	2.58	2.58	2.18	6.50	.52
28.	Ames (Interpo.)	6.58	2.25	6.30	2.26	-2.80	.78
29.	Verhoeff	2.96	1.32	3.15	1.33	4.22	.84

*A negative sign appears in the "critical ratio" column for the four variables (10,11,20,28) which showed a lower level of performance in the retest. Except for the four phoria measurements, all of the other differences indicate improvement in the second series of tests.

Ortho-Rater far stereopsis test shows a difference of only .06, about the same as the real depth tests.

b. Practice effects for monocular depth tests. Three of the four monocular variables show substantial and highly significant gains upon retest: monocular motion parallax (3), AAF "size-constancy" (8), and the Ames size-depth illusion (27). The critical ratios are 5.71, 7.02 and 6.50, respectively. In terms of amount of gain, if the differences between test and retest means are converted into units of area under the normal curve, the three percentages are: .17, .16 and .16. This means that on the retests about 66 per cent of the subjects exceeded the mean of the initial tests for these three variables.

The Ames interposition illusion (28) showed a slight "loss" in the second test period: the critical ratio is 2.80 and the amount of loss in terms of our "standard" unit is .05. This means that the effectiveness of the illusory interposition cues is somewhat reduced at the second trial. The subjects tend to perceive the targets slightly more as they actually are, rather than as the false interposition cues dictate. But the change is slight.

c. Practice effects for the acuity tests. Altogether there are 11 visual acuity variables in the test battery. Table II shows that certain of them exhibit slight gains in the second test series and some of these gains are statistically significant (variables 5, 15, 21). But the gains are small in amount, the largest corresponding to about .08 in "x/sigma x" units. (This means that some 58 per cent of the subjects made retest scores greater than the mean of the initial test distributions.) In the case of two of the Ortho-Rater far acuity tests (variables 10 and 11), there was a slight decline in mean score for the retest series, but the changes were not statistically significant.

In general, the acuity tests show much less practice effect than the depth perception tests.

d. Practice effects for the phoria tests. Only the Ortho-Rater "near lateral" phoria test shows a statistically significant difference between the test and retest means. There is a slight decrease in relative exophoria upon repetition of the test. The amount of difference equals 4 per cent in terms of "normal curve" area, i.e., some 46 per cent of the subjects had retest scores greater than the mean of the first set of scores.

3. Test-Retest Reliability

The last column in Table II presents reliability coefficients for the 28 variables. In accordance with the plan followed in the two preceding sections, these correlations will be discussed briefly for the four types of test variables.

a. Reliability of the binocular depth tests. Taken as a group the "real" depth tests (2, 26, 29) - the two versions of the Howard-Dolman and the Verhoeff - are somewhat more reliable than the "stereoscopic" tests.

The regular Howard-Dolman and the Verhoeff have reliabilities of .79 and .84, while the modified Howard-Dolman with the tilted, colored target has a reliability of only .70.

Among the tests using "simulated" depth, those made with "far vision" are less reliable than the "near stereopsis" tests. The reliabilities of the far stereopsis tests of the three commercial devices are: Sight-Screener, .65, Ortho-Rater, .72 and Telebinocular, .65. The reliabilities for near stereopsis are: Sight-Screener, .79 and Telebinocular, .76.

b. Reliability of the monocular depth tests. Only the interposition test (28) yields a satisfactory level of reliability, with a test-retest correlation of .78. The other three monocular variables all have reliability coefficients below .60. The monocular motion parallax test is the least dependable, with a reliability of .34. The judgments in this situation are unusual and difficult for the subject, and a high degree of variability is not surprising. Careful experimentation will be necessary to establish optimal conditions for the measurement of this aspect of depth discrimination.

Gibson's AAF distance estimation test is more reliable, but the 20-item test shows a test-retest correlation of only .54. Gibson (4, p. 215) reports two "split-half" reliability coefficients for the test, each half consisting of 20 items: the coefficients are .40 and .66, for two methods of scoring ("number right" and a weighted score). Substituting our coefficient of .54 in the Spearman-Brown formula, a reliability of .70 would be predicted for a test of 40 items.

The Ames-Tresselt size-depth test has about the same level of reliability as Gibson's AAF test. The test-retest correlation is .52. This situation is subject to considerable ambiguity; performance in it is probably conditioned by several cues and sets. The subject sees three strips of light, all at the same actual distance, but is told that the middle one is farther or nearer than the other two. He ostensibly "moves" the middle strip backward or forward until it appears to be at the same distance as the other two; actually he changes the height of the variable strip until it seems to be in the plane of the two standard strips. If the "set" changes from one of "tridimensional movement" to "bidimensional matching of height," for many subjects, then unreliability of "scores" would necessarily result. Further experimental work is needed with this situation before its utility in the study of individual differences can be established.

In fact, this last statement can be applied generally to the monocular variables. They probably are highly important determinants of critical aspects of depth perception. But they are also difficult to isolate and control. Careful experimentation must precede the effective use of them as depth perception tests.

c. Reliability of visual acuity tests. For the two Sight-Screener acuity tests the test-retest correlations of .74 and .72 agree rather closely with coefficients reported for this test in the A.G.O. study (.73 and .70 for "both eyes") even though the special slide used in the present study presented Landolt rings instead of letters.

The reliability coefficients secured for the Ortho-Rater acuity tests generally run lower in our study than for the naval population studied by Sulzman, Cook and Bartlett (16), which also supplied the data for part of the A.G.O. study. Our coefficient of .87 for "right-eye" far acuity is somewhat higher than their reliability for this variable. But our near acuity measures particularly show considerably lower test-retest reliability; the coefficients are .59, .70 and .59 (for left, right, and both eyes) whereas coefficients above .80 are reported for all three measures in the earlier study. The reasons for this disparity are not known.

The Telebinocular near acuity variable (21) - the "circles" test with both eyes - has the quite low reliability coefficient of .45. But the two monocular "far acuity" measures show satisfactory reliabilities of .79 and .85 for the right and left eyes, respectively.

d. <u>Reliability of Ortho-Rater phoria tests</u>. The measures of lateral phoria are somewhat more reliable than the two vertical phoria tests. The coefficients for far and near lateral phoria are .76 and .78, while the two vertical phoria tests have reliabilities of .70 and .72. Our vertical phoria measurements are somewhat more reliable than the corresponding variables in the Navy study (.62 and .63 for near and far phoria). But their lateral phoria measurements show higher reliability: .92 and .87 for near and far measures, respectively (17, p. 61).

4. <u>Factor Analysis of the 29 Variables</u>

Two sets of intercorrelations were computed from the test and the retest scores of the 228 subjects. Age was added to the 28 visual variables to make a total of 29. The intercorrelation matrices are shown in Tables III and IV. In each case the correlation coefficients are shown in the upper part of the table and the sixth-factor residuals in the lower.

The centroid factor solution was used and a preliminary set of six factors were extracted from both matrices. The analysis was discontinued at this point, since the standard deviation of the sixth-factor residuals was lower in each case than that of a correlation coefficient of zero. The preliminary factor patterns were then transformed by orthogonal rotation of axes into multiple-factor patterns which conform fairly well to Thurstone's criteria of "simple structure." It will be seen that only the first two of the six factors account for very substantial amounts of the variances of the several tests. Moreover, only three of the factors (I, II and IV) show consistent loadings from test to retest patterns. The factor coefficients for both matrices are shown in Tables V and VI.

a. <u>Factor I. Stereopsis</u>. All of the binocular depth variables (both "real" and "simulated") have significant and substantial loadings on this factor. It is given the name "Stereopsis" because these tests all depend upon variations in stereoscopic angle as a means of securing variations in difficulty of the test items. This factor is presumably the same as that called "Depth Perception" in the AGO study.

TABLE III.

Intercorrelations and Sixth-Factor Residuals for the Initial Scores on the 29 Variables

Intercorrelations*

	Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
	1. Age		-01	-05	-05	-02	-01	07	-13	-06	08	10	08	-07	12	-01	10	02	01	14	06	-09	07	-03	-13	-04	-05	-12	-04	-00
	2. H-Dolman	00		25	47	48	52	40	-02	55	39	38	43	35	34	34	05	-18	13	14	02	26	42	44	32	41	55	01	-13	62
	3. H-D Motion	-02	-02		21	17	24	15	19	18	15	12	12	05	05	10	03	-02	-04	00	01	18	10	19	16	17	13	-01	-02	21
	4. SS F. Stereopsis	01	-04	00		32	68	30	-06	56	23	26	28	26	29	32	08	-14	-01	15	07	26	29	27	38	39	29	-02	-02	46
	5. SS F. Acuity	-06	02	02	01		32	55	-06	44	46	42	56	34	36	40	03	-14	12	10	09	33	46	50	12	32	30	-01	-08	34
	6. SS N. Stereopsis	06	-02	-04	10	-02		36	-01	54	23	29	32	32	45	30	01	-14	02	11	06	31	25	31	41	32	33	-08	-03	63
	7. SS N. Acuity	-04	-05	04	-01	-06	-03		-15	36	38	35	45	45	44	49	12	-12	05	15	18	44	35	34	19	24	18	03	-05	42
	8. AAF-Size	08	-02	10	02	00	-03	-04		-13	-04	-09	-10	-10	-08	-08	-03	13	-01	-06	02	-03	-08	-03	-06	-05	-02	-01	-02	-07
	9. O-R F. Stereo.	-05	-03	05	04	00	-02	04	09		55	43	44	31	36	31	10	-19	08	17	05	30	46	42	39	45	32	-04	00	56
	10. O-R F. Acuity (L)	-06	02	-10	06	03	07	-05	-05	-12		43	57	46	23	34	14	-06	06	17	07	12	27	59	15	31	13	00	-06	28
Residuals	11. O-R F. Acuity (R)	-04	-05	-03	00	-04	01	06	-02	02	-03		56	26	41	30	15	-14	07	24	04	21	58	29	16	29	30	02	04	34
	12. O-R F. Acuity (B)	-02	-01	03	-01	-01	-01	07	02	-02	-08	04		39	43	41	05	-13	04	09	06	33	[illegible]	42	18	31	27	-02	02	40
	13. O-R N. Acuity (L)	05	-05	-06	06	00	03	00	-01	06	13	02	-03		34	43	12	-05	07	11	16	20	21	27	14	26	20	08	-04	48
	14. O-R N. Acuity (R)	06	-01	-03	-02	-01	10	01	-07	-02	06	02	00	02		43	10	-07	02	15	11	30	39	19	22	22	22	06	-04	43
	15. O-R N. Acuity (B)	-01	-04	01	-04	00	01	-02	02	02	00	-02	-01	02	-05		06	-14	06	08	08	37	25	24	32	20	12	02	-10	31
	16. O-R F. V. Phoria	-04	03	02	00	-01	-02	02	03	00	-04	02	00	-02	-01	01		-13	-05	52	00	05	13	-05	01	06	13	06	-03	10
	17. O-R F. L. Phoria	01	07	00	-04	-01	-04	03	-07	-08	02	-03	-07	00	01	02	05		-15	-06	40	-08	-16	-01	-17	-14	-23	04	16	-09
	18. O-R Color	01	00	04	-04	02	-03	02	-04	02	-03	-01	-03	-04	-04	02	04	04		00	-06	09	10	10	08	06	15	04	-12	07
	19. O-R N. V. Phoria	-06	01	-02	-02	-02	-05	02	00	00	-04	-01	03	-05	-02	01	03	00	00		-02	04	18	07	09	08	10	02	01	17
	20. O-R N. L. Phoria	01	05	-02	-01	02	-04	-04	-04	-05	02	-08	-09	-03	-04	00	-01	14	06	04		08	08	05	-08	00	04	01	10	10
	21. TB N. Acuity (B)	05	02	09	-02	-01	-03	05	04	-02	-14	00	-02	-12	00	-02	03	-06	00	00	-07		22	20	23	17	11	12	-14	29
	22. TB F. Acuity (R)	-06	-04	-03	01	01	-05	01	-03	04	08	06	04	03	01	00	-03	-02	00	00	-01	-03		35	12	29	34	01	01	35
	23. TB F. Acuity (L)	03	-01	-08	01	-02	00	-01	-07	-01	11	05	04	01	03	-04	-07	-06	-05	01	-01	-03	-02		13	28	21	-03	-11	29
	24. TB N. Stereo.	01	05	01	04	06	05	-04	-03	-03	00	-07	-05	-04	00	05	-01	01	-01	05	06	-02	-02	-03		45	20	-04	-07	45
	25. TB F. Stereo.	00	-03	01	-03	02	-12	02	03	02	-03	04	04	00	-03	04	-01	-02	00	02	-02	04	03	01	-13		27	-05	-01	38
	26. H-Dolman (Tilt)	-09	09	-04	-06	03	-03	00	-05	-07	02	-03	01	-05	-02	01	-01	00	03	07	10	02	-03	01	03	-02		07	00	41
	27. Ames (Size)	12	-06	04	-01	00	04	-02	03	00	02	-02	00	02	-02	-04	-02	-06	-03	-04	-02	08	-04	02	-04	02	-10		02	-06
	28. Ames (Interpo.)	-08	-04	-02	-01	01	-02	-03	02	03	00	05	05	-01	-03	-01	03	-02	-03	00	-06	03	02	03	-01	03	05	-04		-03
	29. Varhoeff	03	07	-04	-11	-03	02	-04	-02	-01	04	-02	03	-09	02	03	-01	04	00	-01	00	01	-02	01	-02	-05	02	04	-02	

* Decimal points preceding the correlation coefficients have been omitted.

TABLE IV.

Intercorrelations and Sixth-Factor Residuals for the Retest Scores on the 29 Variables

I n t e r c o r r e l a t i o n s *

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
1. Age		-10	-04	-05	-07	-06	03	-01	00	03	12	06	00	12	02	16	02	00	02	02	08	-04	-05	01	00	06	02	-03	-02
2. H-Dolman	01		23	60	41	56	33	00	53	42	43	44	23	26	18	04	00	13	21	09	21	39	46	30	48	41	04	-04	59
3. H-D Motion	-01	-02		26	16	22	12	-08	20	18	10	15	03	08	07	-12	-05	-10	-05	00	08	13	14	25	24	13	01	-17	25
4. SS F. Stereopsis	00	01	-02		44	75	37	-05	69	42	44	42	31	38	28	11	-06	03	27	17	32	31	38	43	53	35	-07	-06	64
5. SS F. Acuity	03	-02	-02	-03		37	62	-12	40	48	47	53	33	40	42	-17	-01	13	17	10	34	40	48	23	33	21	05	-11	37
6. SS N. Stereopsis	01	00	04	-05	-03		40	-03	53	35	40	38	37	46	24	08	-13	05	30	11	31	29	29	48	51	34	-08	-03	72
7. SS N. Acuity	00	00	-01	00	08	00		-14	34	45	48	53	40	50	52	-05	-08	12	29	19	41	35	37	24	28	23	02	-11	38
8. AAF-Size	-04	02	05	00	-04	-03	02		-09	-10	-18	-18	-17	-04	-15	-06	07	-02	-09	-03	00	-07	-12	-12	-12	-08	02	12	-10
9. O-R F. Stereo.	01	04	-03	03	-01	06	00	-01		44	55	48	30	41	26	11	-10	10	22	06	32	42	38	49	62	35	-05	-04	55
10. O-R F. Acuity (L)	00	00	02	02	01	-03	-01	-07	-03		60	68	47	30	42	-02	-02	09	16	14	32	40	62	22	40	28	-05	-10	36
11. O-R F. Acuity (R)	04	01	01	-02	03	00	04	04	03	01		76	34	57	46	08	-06	10	19	05	36	67	40	28	44	34	-04	05	43
12. O-R F. Acuity (B)	00	-01	03	-01	01	-03	01	02	-02	03	05		41	41	48	00	-06	13	14	07	40	59	51	26	38	30	-06	-06	41
13. O-R N. Acuity (L)	-04	-01	-11	01	03	-07	01	01	00	06	00	00		31	41	-05	-02	11	12	14	35	17	25	27	27	20	00	-13	42
14. O-R N. Acuity (R)	-05	-02	02	03	-01	03	-02	04	-06	00	-08	02	01		44	06	-08	06	20	16	38	39	10	26	27	16	-04	-02	42
15. O-R N. Acuity (B)	05	-01	00	-03	00	-04	00	01	00	00	-02	-01	-02	04		00	-14	16	18	04	38	34	25	31	21	10	-02	-07	33
16. O-R F. V. Phoria	02	-07	-05	04	-02	00	-02	-02	-03	02	-08	-04	00	02	01		03	-06	55	03	00	04	-12	04	02	13	-32	08	09
17. O-R F. L. Phoria	-01	-07	-06	-02	-04	02	-04	-01	-07	01	-08	-06	04	08	03	13		-15	-14	38	02	-06	01	-20	-10	-10	-03	13	-06
18. O-R Color	-02	02	09	00	-05	00	-06	03	-04	08	03	04	-01	01	01	01	01		05	-13	18	19	17	06	08	09	05	-11	05
19. O-R N. V. Phoria	-03	03	06	-02	05	00	06	-02	02	-06	07	04	-02	-06	-02	-19	-09	-04		04	07	17	07	17	11	20	01	-02	26
20. O-R N. L. Phoria	-04	-07	-08	03	00	00	-07	03	-04	05	-12	-07	02	06	06	08	16	02	-13		25	01	03	-03	04	06	-07	-02	14
21. TB N. Acuity (B)	03	03	-03	01	04	-01	02	-10	02	-05	-06	-03	02	02	-01	04	00	-12	03	06		30	23	18	24	20	-02	-08	36
22. TB F. Acuity (R)	06	-03	-05	08	-02	-04	-06	-02	00	03	-08	-05	04	01	04	06	06	07	-06	08	00		30	19	30	28	-06	07	37
23. TB F. Acuity (L)	-01	01	-01	-03	03	02	[illegible]	-03	03	-07	04	-01	02	-05	-01	-03	-06	-02	03	-05	00	-08		10	36	28	04	-09	30
24. TB N. Stereo.	00	03	02	-07	-02	03	01	-01	05	-06	07	02	-04	-02	-08	-08	-05	-03	04	-09	-03	-07	08		61	18	-03	-16	54
25. TB F. Stereo.	02	01	-02	-06	-01	02	-02	00	05	-04	08	-02	-05	-04	01	-04	-04	-03	04	-03	-01	-03	03	12		28	-01	-14	50
26. H-Dolman (Tilt)	-06	07	-01	03	-04	-02	-02	-02	03	02	00	02	-02	-07	-09	00	02	02	-01	-04	-02	-01	03	07	04		00	-08	36
27. Ames (Size)	-10	04	01	02	-04	-06	-04	-01	-05	02	-08	00	02	00	-04	06	-01	00	11	03	02	02	00	00	-04	06		-07	04
28. Ames (Interpo.)	02	-02	10	00	-03	05	00	02	-02	-03	-04	00	-05	00	06	-02	-02	-06	00	04	01	01	-01	-03	-01	-04	-01		-11
29. Verhoeff	01	-06	-02	-04	04	-03	05	01	-03	-02	02	01	06	03	00	05	03	00	00	00	01	-04	-01	02	-04	-02	-05	-02	

* Decimal points preceding the correlation coefficients have been omitted.

TABLE V.

Summary of Factor Loadings for Initial Test Matrix

Variables	I Stereopsis	II Retinal Resolution	III Lateral Phoria	IV Vertical Phoria	V ?	VI ?	Commun- ality
1. Age	-.13	.16	.06	-.14	-.18	-.02	.10
2. H-Dolman	.72	.32	.11	-.03	-.12	-.06	.65
3. H-Dolman Motion	.31	.03	.01	.21	.29	.01	.23
4. SS F. Stereopsis	.64	.22	-.13	-.24	.02	.06	.54
5. SS F. Acuity	.36	.56	-.32	-.09	.01	.08	.56
6. SS N. Stereopsis	.70	.22	-.06	-.30	-.08	.09	.64
7. SS N. Acuity	.24	.67	-.18	.11	-.12	-.14	.58
8. AAF-Size	-.19	.00	.01	.15	.21	.02	.10
9. O-R F. Stereo.	.62	.40	-.01	-.12	.07	.14	.59
10. O-R F. Acuity (L)	.31	.53	-.24	.00	.32	.02	.53
11. O-R F. Acuity (R)	.25	.61	.00	-.17	-.24	.07	.52
12. O-R F. Acuity (B)	.27	.67	-.27	-.04	-.02	.19	.63
13. O-R N. Acuity (L)	.27	.49	.20	-.09	-.17	-.07	.40
14. O-R N. Acuity (R)	.25	.55	.13	.01	.09	-.21	.44
15. O-R N. Acuity (B)	.26	.52	-.17	.14	-.24	-.23	.50
16. O-R F. V. Phoria	-.03	.32	.02	.49	.28	-.32	.51
17. O-R F. L. Phoria	.01	-.33	.38	-.21	.15	.23	.36
18. O-R Color	.15	.04	.23	-.06	.13	.09	.11
19. O-R N. V. Phoria	.04	.36	.04	.46	.17	-.32	.48
20. O-R N. L. Phoria	-.14	.26	.37	-.12	.01	.25	.30
21. TB N. Acuity (B)	.29	.36	-.14	.08	-.20	-.21	.32
22. TB F. Acuity (R)	.31	.53	-.01	-.22	-.34	.00	.54
23. TB F. Acuity (L)	.40	.36	.17	.32	.24	.10	.48
24. TB N. Stereo.	.55	.08	-.09	-.21	-.17	-.14	.41
25. TB F. Stereo.	.54	.20	-.03	-.09	.03	.04	.34
26. H-Dolman (Tilt)	.50	.20	-.17	.21	-.18	.00	.39
27. Ames (Size)	-.07	.09	.02	.08	.05	.16	.05
28. Ames (Interpo.)	.01	-.14	.23	.05	.03	.24	.13
29. Verhoeff	.64	.35	-.16	-.22	-.09	.07	.62

TABLE VI.

Summary of Factor Loadings for Retest Matrix

Variables	I Stereopsis	II Retinal Resolution	III ?	IV Vertical Phoria	V ?	VI ?	Communality
1. Age	-.03	.09	-.19	-.12	-.05	-.17	.09
2. H-Dolman	.63	.20	-.36	.00	.04	-.22	.62
3. H-Dolman Motion	.27	.04	-.11	-.20	.02	.12	.14
4. SS F. Stereopsis	.79	.17	-.14	-.09	.22	-.12	.74
5. SS F. Acuity	.32	.61	.13	.14	-.03	-.25	.57
6. SS N. Stereopsis	.79	.13	.02	-.17	-.21	.10	.72
7. SS N. Acuity	.33	.65	-.18	.11	-.22	-.10	.64
8. AAF-Size	-.20	-.06	-.08	.02	.12	-.20	.10
9. O-R F. Stereo.	.72	.25	-.18	.06	.13	.04	.63
10. O-R F. Acuity (L)	.38	.63	-.02	.27	-.21	.05	.66
11. O-R F. Acuity (R)	.45	.65	.04	-.34	.11	-.02	.75
12. O-R F. Acuity (B)	.40	.72	.13	-.18	-.07	-.02	.73
13. O-R N. Acuity (L)	.30	.47	.23	-.09	.04	-.15	.39
14. O-R N. Acuity (R)	.39	.46	-.31	-.01	-.25	.18	.55
15. O-R N. Acuity (B)	.25	.58	-.21	.08	.12	.00	.46
16. O-R F. V. Phoria	.31	-.20	.52	.36	-.26	.03	.60
17. O-R F. L. Phoria	-.01	-.20	-.06	.14	.48	.07	.30
18. O-R Color	.05	.20	.00	.06	.04	.31	.14
19. O-R N. V. Phoria	.43	.03	.25	.31	.23	.11	.42
20. O-R N. L. Phoria	.02	.21	-.09	.07	-.01	.48	.29
21. TB N. Acuity (B)	.25	.50	-.10	.08	-.05	.15	.35
22. TB F. Acuity (R)	.37	.46	.11	-.33	.06	-.29	.51
23. TB F. Acuity (L)	.31	.46	.00	.48	-.04	.16	.57
24. TB N. Stereo.	.63	.05	.09	-.28	.02	-.20	.53
25. TB F. Stereo.	.67	.17	-.23	-.16	-.19	-.09	.59
26. H-Dolman (Tilt)	.45	.17	.09	-.06	.02	.03	.24
27. Ames (Size)	.02	-.13	.29	.07	-.16	-.13	.14
28. Ames (Interpo.)	-.12	-.07	.29	-.05	.16	-.17	.16
29. Verhoeff	.74	.22	.04	-.24	-.12	.13	.68

For the initial test matrix the factor coefficients range from .50 to .72, which means that 25 to 50 per cent of the variance upon these tests is accounted for by Factor I. The retest loadings are slightly higher, on the average; only variable 26 shows a lower coefficient for the retest matrix. Factor I accounts for 20 to 62 per cent of the retest variance of scores on these binocular tests. There are no consistent differences between the real (2, 26, 29) and the simulated depth variables (4, 6, 9, 24, 25) in sizes of factor loadings. The Howard-Dolman has the highest loading for the initial test, but Table VI shows that the two Sight-Screener stereopsis variables have the highest retest loadings. The modified Howard-Dolman test with tilted colored squares as targets has the lowest loading on Factor I among the eight binocular depth tests.

Only one of the four monocular variables has a significant loading on Factor I: monocular motion parallax with the Howard-Dolman rods as targets (3). On both test and retest, approximately 10 per cent of the variance of scores on this variable seems attributable to Factor I. If this factor were a homogeneous determinant involving stereoscopic acuity solely, one would scarcely expect any relationship between it and the monocular variables. But in this instance, with one eye covered, the performances of the subjects apparently are conditioned in part by a factor which is ostensibly "binocular" in nature. Possibly the paradox might involve the concurrent operation of binocular muscular cues associated with convergence.

It is interesting to note that Factor I has significant, although usually low, loadings on all of the visual acuity tests. Furthermore, the coefficients are quite as high for the monocular acuity tests as for the binocular. The percentages of acuity test score variances attributable to Factor I range from approximately 5 to 20 per cent. The manner in which this factor named "Stereopsis" influences monocular acuity test scores cannot be determined from these results. It should be noted that these results differ from those obtained in the AGO study, which found mostly negligible loadings of the "Depth Perception" factor on acuity tests (1, p. 66).

With respect to the phoria tests, the factor coefficients in Table V show no significant loadings for Factor I. But the retest results in Table VI contradict this finding in the case of the two vertical phoria tests. Both the far and the near vertical phoria variables have significant loadings on Factor I; the proportions of test score variance accounted for are approximately 10 and 19 per cent, respectively. Whether or not the disparity in the results for the test and retest series is due to sampling fluctuations is unknown. The AGO study found no significant "Depth" loadings on phoria tests (1, p. 79).

In conclusion, it may be noted that only one of the eight binocular depth perception tests has a significant loading exclusively on Factor I: the Telebinocular near stereopsis test, for the initial test matrix. But although a relatively "pure" measure of Factor I, this variable does not depend upon the factor for a very high proportion of its score variance (30 %). All of the other tests have substantial loadings on Factor II.

b. Factor II. Retinal Resolution. This factor is apparently the same as that identified in the AGO study as the general factor in visual acuity. Inspection of Tables V and VI shows that all of the acuity tests have significant loadings on Factor II. In Table V (initial test), Sight-Screener near acuity (7) and Ortho-Rater far acuity (12) have the highest coefficients for this factor; in both tests about 45 per cent of the variance is accounted for by Factor II. The retest loadings are similar, with the Ortho-Rater variable showing a somewhat higher value (50 per cent of score variance). The Telebinocular acuity tests show substantially lower loadings on Factor II than are found for either the Sight-Screener or the Ortho-Rater. This is in agreement with the AGO study (1, p. 66).

The eight binocular depth variables show mostly low and barely significant loadings on Factor II. The highest loading occurs for Ortho-Rater far stereopsis (9) in Table V; Factor II accounts for 16 per cent of this score variance. In most instances this factor accounts for less than 10 per cent of the variance and several of the factor coefficients are probably not statistically significant. This picture differs somewhat from the results of the AGO study, where the factor named "Retinal Resolution" had substantially higher loadings on all of the depth perception tests than shown in Tables V and VI. The percentages of score variance due to the factor ranged from about 10 to 27 in the AGO study.

In the case of the Ortho-Rater tests, the far acuity variables have considerably higher loadings on Factor II than do the tests of near acuity. But this relationship is reversed for the Sight-Screener, and there are no consistent trends in the loadings for the Telebinocular.

c. Other factors. Tables V and VI show coefficients for four additional factors which were extracted from the two correlational matrices. Except in the case of Factor IV (Vertical Phoria), the patterns for these factors are not consistent from test to retest analysis. It is, therefore, difficult to assign names to them with any degree of assurance. All four of these factors seem to have significant loadings on the one or more of the four phoria tests - both for the test and for the retest series. But there are scattered and inconsistent loadings on several of the other tests, so that it is only w th considerable reservation that any labels at all are assigned to these last four factors in Tables V and VI. This situation probably reflects the lack of enough related variables within the monocular depth and phoria groups to "generate" stable patterns of factorial loadings. The lack of communality means, of course, that most of the test score variance must be provisionally attributed to "specific and error" determinants. With these qualifications, a few comments will be offered concerning the trends exhibited among these four miscellaneous factors.

Factor III. Lateral Phoria (first matrix only). Significant positive loadings on this factor are shown both for the far and near lateral phoria tests in Table V. The percentages of score variance accounted for are approximately 14 in each case. Barely significant negative loadings are found on Factor III for some of the far acuity variables (5, 10 and 12). But none of these relationships shows up clearly in the analysis of the retest matrix.

Table VI shows that for the latter the communalities of the two lateral phoria variables are accounted for mainly by Factors V and VI. And each of these has low, scattered loadings on acuity variables.

Factor IV. Vertical Phoria. This factor has a rather consistent pattern of loadings for both the test and retest analyses. In the initial test series, Factor IV accounts for 23.5 per cent of the score variance of the far vertical phoria test, and 21.5 per cent of the variance of near vertical phoria scores. These proportions go down for the retest series, being 12.8 and 9.8 per cent, respectively. Factor IV also has a low but consistent loading on Telebinocular far acuity (left eye) over the two test series. Otherwise, the loadings on this factor tend to be either insignificant or inconsistent, or both.

d. Specificity of the monocular depth tests. Inspection of Tables V and VI shows that on none of the four monocular depth variables is the score variance determined to any substantial degree by any of the six factors. The test and retest communalities for these four tests are: monocular motion parallax (3), .22 and .14; AAF size-distance estimation (8), .10 and .10; Ames-Tresselt size-depth illusion (27), .05 and .14; Ames-Bakan interposition illusion (28), .13 and .16. The only consistent loading is found in the case of monocular motion parallax, which on the test and the retest has 9.5 and 7.4 per cent of its variance accounted for by Factor I. The anomalous nature of this result has already been discussed.

In general, these four monocular depth tests show very little interrelationship. They seem to be measuring quite specific and mutually independent aspects of "depth perception."

VI. SUMMARY AND CONCLUSIONS

This report presents the results of an analysis of two sets of measurements secured with a battery of 28 visual tests. The primary purpose of the study was to make factorial analyses of these test and retest scores in order to determine the interrelationships among the aspects of visual discrimination involved in "depth perception." From an earlier factorial study reported by the Personnel Research Section of the Adjutant General's Office (1), it seemed probable that tests based upon different cues of depth perception might measure relatively independent functions. The present study, therefore, has included a variety of such situations: monocular and binocular depth perception tests and among the binocular tests, those which involve "real" depth and those utilizing the simulated depth of stereoscopic instruments.

In addition to the factorial analyses, the test and retest data for the several variables were analyzed with reference to three other problems: (a) the form of the distribution of scores for different types of visual discrimination, (b) test-retest reliability, and (c) practice effects.

The principal results of the study can be summarized as follows:

1. Frequency Distributions (See Appendix B)

a. Binocular depth tests. There is no uniformity in the form of the frequency distribution for the eight binocular variables. Marked negative skewness is characteristic of some of them (e.g., the Sight-Screener). But for others the distributions show good differentiation among the subjects and a much closer approximation to normality. By these criteria alone, the Telebinocular test (DC 31-51) is the best of the "simulated" depth tests, and the Verhoeff is the best of the three tests of "real" binocular depth discrimination.

b. Monocular depth tests. The AAF size-distance estimation test differentiates well among the subjects and its distributions approximate about as closely to "normality" as any others in the study. There seem to be no systematic differences between the frequency distributions of the monocular and the binocular depth tests.

c. Visual acuity tests. The marked degree of negative skewness found for acuity tests in the AGO study is found also for certain of the present tests (e.g., the special Sight-Screener Landolt ring test). But the Ortho-Rater acuity tests differentiate much better among the subjects. It seems safe to conclude that negative skewness is not inherently characteristic of the distribution of measurements of visual acuity.

d. Phoria tests. In general, the frequency distributions for all four phoria tests are reasonably satisfactory. There is relatively little skewness and fairly good dispersion of scores.

2. Test-Retest Practice Effects

a. Binocular depth tests. Significant gains in mean scores are found for all eight binocular variables. In every case the difference between test and retest means is more than three times its standard error. In general, the "simulated" depth tests show relatively greater practice effects than the tests of "real" depth perception.

b. Monocular depth tests. Three of the four variables in this group (3, 8, 27) show substantial and highly significant gains upon retest. On the average, some 66 per cent of the subjects at the second testing exceeded the mean of the initial test. The Ames interposition illusion showed a slight "loss" in mean score at the retest.

c. Acuity tests. Some of these tests showed increases in mean score at the repetition of the series, but there were also a few decreases. In general, the acuity tests show much less practice effect than depth perception tests.

d. Phoria tests. "Practice" in the case of these tests can only mean "change" in degree of ocular balance in one or both of the two dimensions. Only the Ortho-Rater near lateral phoria test showed a statistically significant change: a slight decrease in relative exophoria upon repetition of the test.

3. Test-Retest Reliability

a. Binocular depth tests. The tests of real depth (2, 26, 29) have somewhat higher reliability, on the whole, than do the "stereoscopic" machine tests (4, 6, 9, 24, 25). The latter's test-retest correlations range from .65 to .79, while the three "real depth" tests have reliabilities of .79, .84 and .70.

b. Monocular depth tests. Only the Ames interposition test (28) has satisfactory reliability, with a coefficient of .78. The other three variables have reliability coefficients below .60.

c. Acuity tests. The several acuity measures showed a considerable amount of variability in test-retest correlation. Far acuity tests tended to be more reliable than those for near acuity, especially with the Ortho-Rater, (e.g., .72 to .87 for far acuity, and .59 to .70 for near acuity).

d. Phoria tests. The reliability coefficients for the two lateral phoria tests were .76 and .78; the two vertical phoria tests had reliabilities of .70 and .72.

4. Factor Analyses

Two principal factors were identified for both the test and the retest matrices: Stereopsis and Retinal Resolution. The analysis was continued until four additional factors were extracted for each matrix, but only one of these (Vertical Phoria) was clearly identified in both analyses. The factor patterns for the other three were not consistent from test to retest data.

a. Factor I. Stereopsis. All of the eight binocular tests have significant and substantial loadings on this factor; some 25 to 62 per cent of their score variance is accounted for by Factor I. There are no consistent differences in size of factor loading between the "simulated" and the "real" depth variables. Only one of the monocular depth tests, however, has a significant coefficient for Factor I and it accounts for only about 10 per cent of the score variance (monocular motion parallax). All of the visual acuity tests have significant, but mostly low, loadings on Factor I (the percentage of score variance accounted for ranges from 5 to 20). As for the phoria tests, Factor I has no significant loading for the first matrix, but the retest results show significant loadings for the two vertical phoria tests.

b. Factor II. Retinal Resolution. All of the visual acuity tests have significant loadings on this factor. The highest coefficient is found for Ortho-Rater far acuity (retest), where some 50 per cent of the variance of scores is accounted for by Factor II. The eight binocular depth variables have mostly low and barely significant loadings on "Retinal Resolution." And none of the other variables has significant and consistent loadings on this factor, except near lateral phoria. In this case, Factor II accounts for 4.3 and 6.6 per cent of the variance (test and retest).

c. Other factors. The last four of the six factors all have significant loadings for the four phoria tests, either for the initial or for the retest matrix. But only Factor IV, Vertical Phoria, has a fairly consistent pattern for both analyses. The percentage of score variance of the Ortho-Rater vertical phoria test accounted for by Factor IV varies from 9.8 to 23.5.

d. Specificity of the monocular depth tests. These tests showed very low correlation with the other variables, and therefore, had low and mostly insignificant loadings on all of the factors. The only significant factor coefficients were found for monocular motion parallax on Factor I, which accounted for 9.5 and 7.4 per cent of the score variance in the test and retest series.

5. General Conclusions

a. The results of the present study fail to support the view that "depth perception" is a homogeneous behavior category. Judgments of depth and distance based upon different cues or secured in dissimilar situations tend to show only slight interrelationships.

b. Even among tests based primarily upon binocular disparity (stereo acuity), a high degree of non-relationship among scores is found. The factor called "Stereopsis" accounts at most for about 62 per cent of the score variance and on certain of the binocular depth tests the percentage drops as low as 20.

c. Comparisons of "real" depth tests with the "stereoscopic" machine tests fail to show any systematic differences. On the whole the factor loadings on Factor I (Stereopsis) are similar in range and magnitude for the two groups of tests. Moreover, the original intercorrelations are about as high between tests of the two groups as among tests within each category. But again the fact that individual tests are not interchangeable should be stressed; for both groups of tests specific and error variance is fairly high.

d. The monocular tests were all virtually unrelated either among themselves or to other variables included in the study. The lack of communality meant that the factorial composition for these tests could not be determined. Successful and definitive factor analysis in this field must be preceded by extensive experimental research and test development designed to yield a set of several reliable variables for each of the important monocular cues.

e. It is suggested that future investigation might profitably follow either or both of two courses: (1) systematic experimental analysis of the manner in which the several types of visual cues influence judgments of depth and distance, followed by the development of "clusters" of reliable tests based upon each stimulus variable; (2) the direction of experimentation and test development towards the analysis of the visual requirements of a particular type of job. The former approach would establish the basis for a comprehensive factorial analysis of depth perception and related functions. The second approach would yield a more restricted analysis of the organization of visual functions but might provide a set of instruments appropriately weighted with reference to the assessment of visual qualifications for the specific occupational area.

BIBLIOGRAPHY

1. ADJUTANT GENERAL'S OFFICE. Studies in visual acuity. Washington, D. C.: U. S. Government Printing Office, 1948.

2. FONDA, G. E., GREEN, E. L., and HOGAN, F. V., JR. Comparison of results of the Sight-Screener and clinical tests. U.S.A.F., Sch. Aviat. Medicine, Project No. 48[illegible], Report No. 1, Sept. 1946.

3. FRY, G. A. Measurement of the threshold of stereopsis. Optom. Weekly, 1942, 33, 1029-103[illegible].

4. GIBSON, J. J. (Ed.) Motion picture testing and research. AAF Aviation Psychology Program Research Reports, Report No. 7. Washington, D. C.: U. S. Government Printing Office, 1947.

5. GIBSON, J. J. Studying perceptual phenomena. Ch. 6 in T. G. Andrews (Ed.). Methods of Psychology. New York: Wiley, 1948. Pp. 158-188.

6. GRAHAM, C. H. Visual space perception. Federation Proceedings, 1943, 2, 115-122.

7. GRAHAM, C. H., BAKER, KATHERINE E., HECHT, MARESSA, and LLOYD, V. V. Factors influencing thresholds for monocular movement parallax. J. exp. Psychol., 1948, 38, 205-223.

8. HIRSCH, M. J., and WEYMOUTH, F. W. Distance discrimination. V. Effect of motion discrimination and distance on monocular and binocular distance discrimination. J. aviat. Med., 1947, 18, 594-600.

9. HIRSCH, M. J., and WEYMOUTH, F. W. Distance discrimination VI. The relationship of visual acuity to distance discrimination. J. aviat. Med., 1948, 19, 56-58.

10. HOLWAY, A. H., and BORING, E. G. Determinants of apparent visual size with distance variant. Amer. J. Psychol., 1941, 54, 21-37.

11. IMUS, H. A. Comparison of the Ortho-Rater with clinical ophthalmic examinations. J. aviat. Med. 1949, 20, 2-23.

12. KIRSCHBERG, L. S. S. Depth perception and flying ability. Arch. Ophthal. 1946, 36, 155-170.

13. MOFFIE, D. J., BARKLEY, K. L., OLSON, H. C., MASSEY, B. T., and CLINTON, C. L. Progress report no. 2 - occupational vision research project at North Carolina State College. N. Car. Optometrist, 1949, 13, Nos. 5-6, 3-19.

14. ROWLAND, W. M., and LOUISE S. A comparison of three tests of depth perception. U.S.A.F., Sch. Aviat. Medicine, Project No. 238, Report No. 1.

15. SLOANE, A. E., and GALLAGHER, J. R. Stereopsis: a comparison of the Howard-Dolman and Verhoeff test. Arch. Ophthal., 1945, 34, 357-359.

16. SULZMAN, J. H., COOK, E. B., and BARTLETT, N. R. The reliability of visual acuity scores yielded by three commercial devices. J. appl. Psychol., 1947, 31, 236-240.

17. SULZMAN, J. H., COOK, E. B., and BARTLETT, N. R. The validity and reliability of heterophoria scores yielded by three commercial optical devices. J. appl. Psychol., 1948, 32, 56-62.

18. VERHOEFF, F. H. Simple quantitative test for acuity and reliability of binocular stereopsis. Arch. Ophthal., 1942, 28, 1000-1019.

19. ZEGERS, R. T. Monocular movement parallax thresholds as functions of field size, field position, and speed of stimulus movement. J. Psychol., 1948, 26, 447-498.

APPENDIX A

Descriptions of the Variables used in the Study

1. Age

Originally it was planned to restrict the subjects (all male) to ages 18 - 25. But when it became difficult to secure such subjects for test and retest sessions, it was decided to remove the age limits and to include age as a variable in the correlation matrix. The ages ranged from 17 to 39, but most of the subjects fell within the range 18 - 25 years (see the frequency distribution in Appendix B).

2. Howard-Dolman Binocular Depth Test

This familiar test presents two black vertical rods, the right-hand one at a fixed distance from the subject and the other adjustable in depth by means of cords held by S. The model used was secured from C. H. Stoelting. A 20-watt fluorescent lamp was mounted longitudinally at the top of the frame so as to provide uniform illumination, and a cardboard screen at the front concealed all of the instrument except the aperture through which S viewed the rods. S viewed the rods at a distance of 20 feet from the fixed rod. The position of his head was kept constant by means of a fixed head rest. S made 10 settings of the variable rod, being instructed to adjust it to the same distance as the fixed rod. Each trial began with the variable rod in a different (but predetermined) distance from the standard one, as follows: +190, -10, -90, +60, 0, -130, -50, +100, +150, -180 millimeters. After each trial S dropped the cords so as to avoid judgments based upon spurious non-visual cues. The average error of the 10 trials in millimeters (irrespective of signs) is S's score.

3. Monocular Motion Parallax (with Howard-Dolman apparatus)

Because a precision instrument was not completed in time for use in the study, a procedure was devised in which S judged the relative distance of the Howard-Dolman rods while his head was in motion. He rested his chin on a carriage which he moved back and forth on a small track perpendicular to the line of regard. The distance covered by the eye in a single excursion represented an angle of 4 degrees, at a distance of 10 feet from the fixed rod. The rods were viewed with the "preferred" eye, while the other eye was shielded by an opaque goggle. E set the variable rod each time, and asked S to judge whether it was farther or nearer than the standard after moving his head through two complete "cycles" along the track (right, left, right, left) in synchronism with a metronome which beat once in two seconds. The series began with the variable rod 10 millimeters from the standard, and four trials were made at that distance (two "nearer" and two "farther" from the standard.) (It was determined in preliminary experimentation that subjects could not discriminate separations of less than 10 mm.). In successive series the distance was increased by increments of 10 mm. until S had responded correctly three times in succession for each of the four columns of trials just noted. S's score was recorded as the mean separation of the rods at which S had judged correctly <u>twice</u> in succession. A sample record will illustrate the procedure:

Distance in mm.	Variable Rod Farther 1st trial	Variable Rod Farther 2nd trial	Variable Rod Nearer 1st trial	Variable Rod Nearer 2nd trial
10	nearer	farther	nearer	farther
20	nearer	nearer	farther	farther
30	farther	nearer	nearer	nearer
40	farther	farther	nearer	nearer
50	farther	farther	nearer	nearer
60		farther		

The mean of the distances corresponding to the four responses underlined (the second consecutive correct response in each column) is 42.5 mm. This would be the "score" for such a record.

4, 5, 6, 7. Sight-Screener Tests

Instead of the standard targets, a special slide prepared according to specifications of Dr. Louise Sloan was used. Tests for only four functions were made: far and near stereopsis, and far and near acuity.[5]

Stereopsis Targets (4, 6). The test is essentially the same as in the "standard" slide, except that there are three items in the practice series and at each of the four levels of difficulty. Thus there is a total of 15 items instead of the standard five. The Shepard-Fry per cent stereopsis values of the "practice" (first) series are 45%, 60% and 75%; then three trials each are made in succession at 45%, 60%, 75% and 105%.[6] In each item one of the five circles is presented to the two eyes so as to produce binocular parallax and hence tends to "stand out" from the other circles.

Acuity Targets (5, 7). The near acuity test is made at 14 inches, whereas the far test is at the optical equivalent of 20 feet. Landolt rings are used instead of the letters of the standard slide. Moreover, the sizes of the rings cover a range from 20/40 to 20/10 Snellen units, whereas the standard range is from 20/200 to 20/10. But there are only nine levels in the latter whereas the special slide developed for Dr. Sloan has levels as follows: 20/40, 20/30, 20/25, 20/23, 20/20, 20/18, 20/17, 20/15, 20/13, 20/10. There are five items for each level except 20/20, which has 10. The subject is required to discriminate the position of the break in the ring on all items in order to receive a score for the level in question. The maximum score is 55.

---

[5] The "Sight-Screener" and the special test slide were kindly provided by the American Optical Co., through the courtesy of Mr. Paul K. Fryer.

[6] See Fry, (3) for the description of this unit based upon the stereoptic angle.

8. AAF Distance Estimation Test CP212-A.

This is the photographic test referred to above in discussing Gibson's work involving "size-constancy" as a basis of depth perception (4). There are 20 photographs of a flat, uniform terrain. Each picture shows a set of 15 comparison stakes which were arranged in a row 14 yards in front of the camera. The comparison stakes were set in an arc, 20 inches apart, except for a gap of eight feet in the center of the arc. (The subject observes the target stake through this gap.) The "standard" stakes varied from 27 to 83 inches, in 4-inch steps. The width of both standard and target stakes varied irregularly between two and four inches.

The four target stakes were 63, 67, 71 and 75 inches in height and these were set at 28, 56, 112, 224 and 448 yards. Thus there were 20 photographs, each target appearing at each of the five distances. A random order of presentation was used. The photographs were placed directly in front of the subject upon a neutral gray background. S sat about 15 inches from the photograph, which was evenly illuminated by two shielded 15-watt fluorescent bulbs. S simply matched the target to one of the standard stakes in terms of height. His responses were scored as follows: correct match, 4; error by one stake, 3; error by two stakes, 2; error by three stakes, 1; error of four or more stakes, 0. The maximum score is 80. This scoring method proved to yield higher reliability in the preliminary studies than did either Gibson's weighted score or the number correct.

9 - 20 (inclusive). Ortho-Rater Tests

This instrument is manufactured by the Bausch and Lomb Optical Co. The entire series of tests was administered, although originally it was planned to use only the depth and acuity tests. A considerable number of subjects were tested before it was decided to include the remainder of the series (phoria and color tests) in the battery. The order of the tests differs, therefore, from that prescribed in the Bausch and Lomb manual. The order given in Table I was followed for all subjects, in both the initial and the retest sessions. Otherwise, the standard conditions for this test were maintained. The subtests and the range of scores for each are as follows:

9. F-6, Far stereopsis: scores, 1 - 9.
10. F-5, Far acuity (left eye): scores, 1 - 15 (20/200 to 20/13)
11. F-4, Far acuity (right eye): scores, 1 - 15 (20/200 to 20/13)
12. F-3, Far acuity (both eyes): scores, 1 - 15 (20/200 to 20/13)
13. N-3, Near acuity (left eye): scores, 1 - 15 (20/200 to 20/13)
14. N-2, Near acuity (right eye): scores, 1 - 15 (20/200 to 20/13)
15. N-1, Near acuity (both eyes): scores, 1 - 15 (20/200 to 20/13)
16. F-1, Far vertical phoria: scores, 1 - 9
17. F-2, Far lateral phoria: scores, 1 - 15
18. F-7, Color: scores, 0 - 6
19. N-4, Near vertical phoria: scores, 1 - 9
20. N-5, Near lateral phoria: scores, 1 - 15

21 - 25. Keystone Telebinocular Test

The three Telebinocular acuity tests listed in Table I are included in the standard series of slides distributed with this instrument (Keystone Visual Survey Series). In order of administration they are as follows: (a) Test XIV which tests near vision for both eyes by having subject indicate whether a series of small circles are filled with lines, dots or solid gray (score, 0 - 22); (b) Test V (far acuity, right eye) requires S to locate a dot in a series of receding signboards in one of four positions (score 0 - 10); (c) Test VI (far acuity, left eye) is the same as V.

The Telebinocular depth test consists of a special set of stereoscopic slides called the "Aviators Unit. DC 31-53". There are 23 cards in the entire series, with Shepard-Fry "per cent stereopsis" values ranging from 1% to 110%. Each card shows three vertical columns of test objects (letters and numbers), and one object appears to stand out in front of the remainder in each column. The 12 odd-numbered cards were used in the test series, making a total of 36 items in the test. Slides 32, 34 and 36 were used as practice slides. Two scores are available: the "per cent stereopsis" value of the last slide for which all responses are correct and the total number of items correct by the same standard.

26. Howard-Dolman-Weinstein[7]

This adaptation of the Howard-Dolman device was designed primarily to introduce two types of variation into this "real depth" situation: (a) both the fixed and the movable rods were attached to their bases by universal joints, so they could be tilted to any desired angle; (b) 3" x 3" colored squares were attached to the tops of the rods. In this study the fixed (right hand) rod was in the vertical position and its square was light green. The variable rod was tilted 30 degrees to the left of the vertical position, with the result that its orange square faced the subject with two of its corners in the vertical plane. The distance between the centers of the two squares was 12 inches, and the exposed length of each rod-square unit was 11½ inches. The subject viewed the units from a distance of 10 feet. The procedure was the same as with the regular Howard-Dolman (see descriptions for variable 2 above), including a headrest for S.

27. Ames-Tresselt Size-Depth Illusion[8]

This test situation represents an adaptation of one of the Ames visual demonstrations. A device was constructed by which three vertical bars of light were viewed by S at a distance of 11 feet in a completely darkened booth. S was given the following instructions: "At the end of this alley

[7] Mr. Sidney Weinstein, one of the research assistants on the project, developed this instrument.

[8] Dr. Margaret Tresselt was responsible for the development and standardization of this test, in accordance with a blueprint secured from the Hanover Institute.

there are three illuminated rods. The ones at the sides will always be at the same distance from you, but the center rod will be either nearer to you or farther from you. The device before you controls the forward and backward motion of the center rod. By turning it to the right you can move the center rod further away. By turning to the left you can bring it closer. The object of this test is for you to line up all three rods in the same plane, at the same distance from you. I will adjust the center rod to a new position each time. You are to line them up each time, tell me when you finish, and then close the screen door. I will tell you when to open it again."

Actually, all three strips of light were in the same plane and the center one varied only in height. Three heights of the outer "fixed" lights were used: 3, $4\frac{1}{2}$ and 6 inches. For each "standard" height the experimenter set the variable center light, in successive trials, at five different heights, making a total of 15 trials for the entire test. After each trial S pulled a screen across his line of vision so the changes in the stimuli could be made without his observing them.

28. Ames-Bakan Interposition Illusion[9]

Another Ames demonstration served as the basis for this "test." The cue of "interposition" is involved in a series of colored cards arranged at different distances from the subject so as to appear to overlap each other. Three sets of three cards each were used and in each set the subject was asked to name the colors of the squares in order of distance from him. Each set of squares was viewed through a separate cylindrical metal tube, $16\frac{1}{2}$ inches long and eight cm. in diameter. The tubes were entirely open at the end which projected into the enclosed booth containing the test squares, but closed at the other end except for an aperture of about 3 mm. through which S viewed the three targets. The arrangement of the targets and the lamps followed the blueprint secured from the Hanover Institute.

The object of the "test" is to determine S's degree of success in using interposition as a cue of distance. But the use of the cue means that the phenomenal order of the squares in space is the reverse of the physical order for the second and third sets of targets. In the first set the two orders are the same. Thus there are nine judgments to be made, and the score is the number "right" by phenomenological consensus.

29. Verhoeff Stereopter

This instrument, widely used as a clinical test of stereopsis, is manufactured by the American Optical Company. A detailed description is given in Verhoeff's paper (18). The device presents three black vertical

---

9 Mr. Paul Bakan, a research assistant on the project, developed and standardized this test, in accordance with a blueprint secured from the Hanover Institute.

strips simultaneously in an illuminated rectangular window. One of the three targets lies in a plane either in front of or behind the other two. Four sets of targets are available, and by rotating the device 180 degrees these can be repeated, in a different spatial arrangement, to give a total of eight items. The object of the test is to find the maximum distance at which S makes eight correct judgments. Obviously, the closer the objects the greater the degree of binocular parallax and the lower the stereoscopic acuity of the subject. The author recommends a score in terms of "Snellen" notation, in which a distance of 100 cm. is taken as a standard. For the present study a simple numerical scoring method is used: one unit for each of the levels of difficulty at which all eight items are correct. The following are the distances, the corresponding Snellen-type values, and the numerical scores used in this study:

Target Distance	"Snellen Score"	Unit Score
200 cm.	20/10	6
150 cm.	20/15	5
100 cm.	20/20	4
67 cm.	20/30	3
50 cm.	20/40	2
40 cm.	20/50	1
Less than 40 cm.	Above 20/50	0

It is essential that head movement and movement of the target be eliminated, and also that distance be measured accurately. Neither of these controls prevails as the test is customarily used. Moreover, the illumination of the targets by means of a flashlight bulb may result in variable light intensity as the battery ages. In the present study, three steps were taken to control these conditions: (1) the instrument was mounted on a heavy steel bar along which it could be set at the specified distances; (2) S's head was held in a padded headrest; (3) a 6-8 volt bulb operated from a transformer replaced the flashlight bulb.

APPENDIX B

Frequency Distributions for the 29 Variables

TABLE B-1. Frequency Distributions of Ages of the Subjects

Age	Frequency
Above 30	5
30	1
29	4
28	4
27	12
26	12
25	25
24	31
23	24
22	29
21	28
20	22
19	18
18	12
17	1
	228
Mean	21.86
S. D.	2.96

TABLE B-2. Frequency Distributions for the Howard-Dolman Test

Score	Test	Retest
Above 97	15	10
89-97	5	4
80-88	4	4
71-79	10	6
62-70	7	12
53-61	11	8
44-52	15	6
35-43	10	21
26-34	30	27
17-25	55	47
8-16	58	76
Below 8	8	7
	228	228
Mean	35.31	31.85
S. D.	27.72	25.56

TABLE B-3. Frequency Distributions for the H-D Motion Parallax Test

Score	Test	Retest
Above 52	31	19
50-52	10	2
47-49	8	4
44-46	9	6
41-43	8	6
38-40	25	17
35-37	24	12
32-34	21	24
29-31	20	18
26-28	21	25
23-25	33	75
Below 22	18	20
	228	228
Mean	35.87	31.42
S. D.	10.73	9.75

TABLE B-4. Frequency Distributions for the Sight-Screener Far Stereopsis

Score	Test	Retest
15	90	116
14	23	38
13	23	14
12	12	11
11	10	8
10	10	4
9	7	4
8	8	5
7	7	6
6	6	6
5	7	1
4	2	3
3	7	1
2	8	3
1	6	1
0	2	2
	228	228
Mean	11.48	12.78
S. D.	4.43	3.58

TABLE B-9. Frequency Distributions for Ortho-Rater F-6, Far Stereopsis

Score	Test	Retest
8	38	59
7	34	39
6	40	32
5	32	28
4	29	16
3	21	26
2	18	11
1	13	11
0	3	6
	228	228
Mean	5.13	5.49
S. D.	2.17	2.31

TABLE B-10. Frequency Distributions for Ortho-Rater F-5, Far Acuity (Left)

Score	Test	Retest
15 (20/13)	6	7
14 (20/14)	2	2
13 (20/15)	14	14
12 (20/17)	30	38
11 (20/18)	48	43
10 (20/20)	30	23
9 (20/22)	33	29
8 (20/25)	30	30
7 (20/29)	16	16
6 (20/33)	9	10
5 (20/40)	2	7
4 (20/50)	3	5
3 (20/67)	3	2
2 (20/100)	2	2
	228	228
Mean	9.76	9.70
S. D.	2.43	2.61

TABLE B-11. Frequency Distributions for Ortho-Rater F-4, Far Acuity (Right)

Score	Test	Retest
15 (20/13)	2	5
14 (20/14)	6	3
13 (20/15)	19	12
12 (20/17)	40	52
11 (20/18)	51	42
10 (20/20)	17	15
9 (20/22)	34	39
8 (20/25)	26	23
7 (20/29)	10	19
6 (20/33)	12	7
5 (20/40)	1	1
4 (20/50)	3	3
3 (20/67)	1	3
2 (20/100)	3	2
1 (20/200 and above)	3	2
	228	228
Mean	9.90	9.87
S. D.	2.62	2.58

TABLE B-12. Frequency Distributions for Ortho-Rater F-3, Far Acuity (Both)

Score	Test	Retest
15 (20/13)	0	2
14 (20/14)	4	6
13 (20/15)	26	28
12 (20/17)	54	62
11 (20/18)	47	43
10 (20/20)	35	27
9 (20/22)	21	19
8 (20/25)	23	21
7 (20/29)	9	11
6 (20/33)	3	3
5 (20/40)	4	3
4 (20/50)	1	1
3 (20/67)	1	2
	228	228
Mean	10.48	10.63
S. D.	2.04	2.16

TABLE B-13. Frequency Distributions for Ortho-Rater N-3, Near Acuity (Left)

Score	Test	Retest
15 (20/13)	3	3
14 (20/14)	11	5
13 (20/15)	17	17
12 (20/17)	35	43
11 (20/18)	55	57
10 (20/20)	30	31
9 (20/22)	37	39
8 (20/25)	21	23
7 (20/29)	10	4
6 (20/33)	4	2
5 (20/40)	0	0
4 (20/50)	0	1
3 (20/67)	0	0
2 (20/100)	2	2
1 (20/200 and above)	1	1
	228	228
Mean	10.28	10.41
S. D.	2.38	2.02

TABLE B-14. Frequency Distributions for Ortho-Rater N-2, Near Acuity (Right)

Score	Test	Retest
15 (20/13)	4	0
14 (20/14)	6	3
13 (20/15)	18	26
12 (20/17)	30	41
11 (20/18)	55	57
10 (20/20)	43	39
9 (20/22)	45	35
8 (20/25)	13	11
7 (20/29)	2	9
6 (20/33)	4	1
5 (20/40)	1	3
4 (20/50)	2	1
3 (20/67)	0	1
2 (20/100)	2	1
1 (20/200 and above)	3	2
	228	228
Mean	10.26	10.47
S. D.	2.27	1.89

TABLE B-15. Frequency Distributions for Ortho-Rater, N-1, Near Acuity (Both)

Score	Test	Retest
15 (20/13)	1	2
14 (20/14)	3	6
13 (20/15)	15	12
12 (20/17)	40	44
11 (20/18)	47	62
10 (20/20)	61	51
9 (20/22)	38	32
8 (20/25)	16	13
7 (20/29)	3	2
6 (20/33)	2	3
5 (20/40)	0	0
4 (20/50)	1	0
3 (20/67)	0	1
2 (20/100)	1	0
	228	228
Mean	10.39	10.60
S. D.	1.65	1.64

TABLE B-16. Frequency Distributions for Ortho-Rater F-1, Far Vertical Phoria

Score	Test	Retest
Above 9	5	3
9	1	4
8	12	14
7	71	62
6	97	109
5	29	24
4	6	7
3	5	2
2	1	2
1	1	1
	228	228
Mean	6.24	6.25
S. D.	1.21	1.17

TABLE B-17. Frequency Distributions for Ortho-Rater F-2, Far Lateral Phoria

Score	Test	Retest
Above 14	3	4
14	1	0
13	1	3
12	4	0
11	6	9
10	25	13
9	32	35
8	35	55
7	40	31
6	25	32
5	22	16
4	12	8
3	6	6
2	7	5
1	1	4
0	8	7
	228	228
Mean	7.13	7.18
S. D.	2.79	2.72

TABLE B-18. Frequency Distributions for Ortho-Rater F-7, Color

Score	Test	Retest
6	51	49
5	110	125
4	46	37
3	13	9
2	8	8
	228	228
Mean	4.80	4.87
S. D.	.97	.91

TABLE B-19. Frequency Distributions for Ortho-Rater N-4, Near Vertical Phoria

Score	Test	Retest
Above 9	3	2
9	0	1
8	1	5
7	10	4
6	43	52
5	92	91
4	61	49
3	10	14
2	7	9
1	1	1
	228	228
Mean	4.89	4.92
S. D.	1.23	1.29

TABLE B-20. Frequency Distributions for Ortho-Rater N-5, Near Lateral Phoria

Score	Test	Retest
Above 14	6	7
14	5	3
13	4	5
12	5	8
11	11	4
10	16	14
9	22	18
8	24	18
7	24	24
6	31	33
5	29	33
4	19	18
3	9	21
2	5	6
1	5	7
0	13	9
	228	228
Mean	6.84	6.50
S. D.	3.46	3.44

TABLE B-21. Frequency Distributions for Telebinocular Near Acuity (Both)

Score*	Test	Retest
22 (105%)	14	16
20 (103%)	26	31
18 (102%)	33	53
16 (100%)	68	56
14 (90%)	32	36
12 (80%)	41	25
10 (70%)	1	1
9 (60%)	7	6
7 (50%)	6	3
Below 7	0	1
	228	228
Mean	15.57	16.21
S. D.	3.34	3.27

TABLE B-22. Frequency Distributions for Telebinocular Far Acuity (Right)

Score	Test	Retest
10 (105%)	103	114
9 (103%)	31	33
8 (102%)	22	17
7 (100%)	10	12
6 (98%)	8	9
5 (96%)	11	12
4 (92%)	8	6
3 (84%)	16	9
2 (65%)	11	10
1 (50%)	2	2
0	6	4
	228	228
Mean	7.73	8.09
S. D.	2.94	2.71

TABLE B-23. Frequency Distributions for Telebinocular Far Acuity (Left)

Score	Test	Retest
10 (105%)	115	109
9 (103%)	41	56
8 (102%)	5	2
7 (100%)	14	12
6 (98%)	13	5
5 (96%)	8	15
4 (92%)	10	8
3 (84%)	7	7
2 (65%)	7	7
1 (50%)	4	3
0	4	4
	228	228
Mean	8.13	8.20
S. D.	2.71	2.65

TABLE B-24. Frequency Distributions for Telebinocular Near Stereopsis

Score	Test	Retest
33 (100%)	0	1
30 (90%)	1	2
27 (80%)	11	21
24 (70%)	32	44
21 (60%)	26	26
18 (50%)	58	55
15 (40%)	13	15
12 (30%)	45	35
9 (20%)	17	11
6 (10%)	10	8
3 (1%)	5	5
0	10	5
	228	228
Mean	16.00	17.80
S. D.	6.76	6.70

*The uneven intervals in this distribution table reflects the fact that at certain levels of difficulty only one item is presented, whereas the remaining levels have two items each. To make the scores comparable with other results of the test, each "per cent" score and its corresponding item count are given. The scale represents actual scores, not class intervals.

TABLE B-25. Frequency Distributions for Telebinocular Far Stereopsis

Score	Test	Retest
33 (100%)	0	0
30 (90%)	1	4
27 (80%)	12	19
24 (70%)	16	29
21 (60%)	19	20
18 (50%)	46	52
15 (40%)	28	26
12 (30%)	60	49
9 (20%)	27	13
6 (10%)	4	6
3 (1%)	9	7
0 (0%)	6	3
	228	228
Mean	14.91	16.83
S. D.	6.22	6.45

TABLE B-26. Frequency Distributions for the Howard-Dolman-Weinstein Test (tilted target, at 10 feet)

Score	Test	Retest
Above 87	7	3
80-87	0	0
72-79	2	3
64-71	9	3
56-63	8	5
48-55	9	13
40-47	16	16
32-39	27	20
24-31	24	29
16-23	45	38
8-15	60	72
0-7	21	26
	228	228
Mean	27.61	24.27
S. D.	20.64	18.19

TABLE B-27. Frequency Distributions for the Ames Size-Depth Illusion

Score	Test	Retest
Above 9.0	18	11
8.3-9.0	3	1
7.5-8.2	5	6
6.7-7.4	8	3
5.9-6.6	10	2
5.1-5.8	8	3
4.3-5.0	12	4
3.5-4.2	19	4
2.7-3.4	30	30
1.9-2.6	38	47
1.1-1.8	52	90
Below 1.1	25	27
	228	228
Mean	3.49	2.58
S. D.	2.58	2.18

TABLE B-28. Frequency Distributions for the Ames- Bakan Interposition Illusion

Score	Test	Retest
9	91	87
8	0	0
7	30	22
6	19	17
5	27	37
4	41	44
3	19	21
2	1	0
	228	228
Mean	6.58	6.30
S. D.	2.25	2.26

TABLE B-29. Frequency Distributions for the Verhoeff Stereopter Test

Score	Test	Retest
6 (20/10 - 200 cm.)	1	3
5 (20/15 - 150 cm.)	10	14
4 (20/20 - 100 cm.)	75	89
3 (20/30 - 67 cm.)	91	80
2 (20/40 - 50 cm.)	21	13
1 (20/50 - 40 cm.)	4	9
0 (Above 20/50)	26	20
	228	228
Mean	2.96	2.25
S. D.	3.15	1.33

290350

28/3.4 STI-ATI-205 308 * UNCLASSIFIED

Psychological Research Center, New York U., N.Y.

A FACTORIAL ANALYSIS OF DEPTH PERCEPTION TESTS, by Lyle H. Lanier and Margaret E. Tresselt. (1950) 42p incl tables.

USA Contr. No. WSW-2496. AGO - Rept 831

AD-A800 018

DIV: Psychology & Human Engineering (28)

SECT: Physiological Psychology (3)

Psychometrics (4)

SUBJECT HEADINGS

Depth perception - Measurement

Visual acuity

VISUAL ACUITY

DEPTH

NTIS, Auth: ARI ltr, 13 Nov 79

(Copies obtainable from ASTIA-DSC)

MICROFILMED

25 Depth Perception

UNCLASSIFIED

28/3,4 STI-ATI-205 308 UNCLASSIFIED

Psychological Research Center, New York U., N.Y.

A FACTORIAL ANALYSIS OF DEPTH PERCEPTION TESTS, by Lyle H. Lanier and Margaret E. Tresselt. (1950) 42p incl tables. USA Contr. No. WSW-2496.

DIV: Psychology & Human Engineering (28)

SECT: Physiological Psychology (3)

Psychometrics (4)

SUBJECT HEADINGS

Depth perception - Measurement

Visual acuity

(Copies obtainable from ASTIA-DSC)

MICROFILMED

UNCLASSIFIED